TECHNICAL REPORT 79-5
MARCH, 1979

LEVEL II

B.S. 12

AN IMPROVED MODEL AND COMPUTER PROGRAM FOR BAYESIAN M-GROUP REGRESSION

BY

IVO W. MOLENAAR AND CHARLES LEWIS
RIJKSUNIVERSITEIT GRONINGEN

REPORT PREPARED UNDER OFFICE OF NAVAL RESEARCH
CONTRACT #00014-77-C-0428
MELVIN R. NOVICK, PRINCIPAL INVESTIGATOR
UNIVERSITY OF IOWA
IOWA CITY, IOWA

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: Technical Report 79-5 ✓	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER: (14) TR-79-5	
4. TITLE (and Subtitle): (6) AN IMPROVED MODEL AND COMPUTER PROGRAM FOR BAYESIAN M-GROUP REGRESSION	5. TYPE OF REPORT & PERIOD COVERED: Research Report 10/78-3/79
	6. PERFORMING ORG. REPORT NUMBER: Technical Report 79-5
7. AUTHOR(s): (10) Ivo W. Molenaar and Charles Lewis	8. CONTRACT OR GRANT NUMBER(s): (15) N00014-77-C-0428
9. PERFORMING ORGANIZATION NAME AND ADDRESS: University of Iowa, Division of Educational Psychology ✓, Iowa City, Iowa 52242	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 6115 3N; RR 042-04; RR042-04-01; NR 150-404
11. CONTROLLING OFFICE NAME AND ADDRESS: Personnel and Training Research Programs, Office of Naval Research (Code 458), Arlington, Virginia 22217 (11)	12. REPORT DATE: Mar[redacted]79 (12) 63
	13. NUMBER OF PAGES: 49
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): (17) RR042041 (16) RR04204	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release, distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

(9) Research rept. Oct 78 - Mar 79,

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Bayesian linear model analysis, m-group regression, convergence

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This paper presents a modified model and computer algorithm for the estimation of regressions in m-groups. The major features of the new approach are: 1) The use of parameters with constant values for slopes across groups whenever the prior or the data indicate that this is desirable, 2) the use of a residual variance constant across groups, 3) independent priors for regression parameters, 4) scales of the predictors at the grand mean rather than

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE S/N 0102-LF-014-6601

Unclassified 411 417

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. Abstract, Continued

at the so-called "Ideal scaling points", 5) transformation of all variables including the criterion to mean zero and variance one at the start of the calculations, 6) an improved search procedure for finding the mode of the posterior distribution. This new solution provides a greatly speeded-up algorithm which defines the posterior mode precisely.

AN IMPROVED MODEL AND COMPUTER PROGRAM FOR BAYESIAN M-GROUP REGRESSION

Ivo W. Molenaar and Charles Lewis
University of Groningen

AN IMPROVED MODEL AND COMPUTER PROGRAM FOR BAYESIAN M-GROUP REGRESSION

Ivo W. Molenaar and Charles Lewis*
FSW, Rijksuniversiteit Groningen (Neth.)

1. Summary

When multiple regression equations are to be estimated for m groups which are supposed to be comparable though not identical, both the pooled estimates and m separate least squares estimates per group may be suboptimal. Lindley, Novick, Jackson and others have advocated a Bayesian estimation procedure in which the estimates would be weighted averages of the separate estimates per group on one hand and some pooled estimate on the other hand, with weights determined essentially by the data. This extension of the Kelley formula for regression to the mean has proven its value in several cross-validation studies (Novick, Jackson, Thayer & Cole, 1972; Lissitz and Schoenfeldt, 1974; Shigemasu, 1976; Jansen, 1977). The modal posterior values for intercepts, slopes and residual variances, however, are not easy to obtain. The procedure outlined by Novick et al. (1972) and Jones and Novick (1972) still poses some numerical and methodological problems. The present paper presents a modified algorithm removing most of the deficiencies. It remains true,

*Supported in part under ONR Contract #N00014-77-C-0428, Melvin R. Novick, principal investigator. Opinions stated herein are those of the authors and not those of the supporting agency.

however, that m-group regression is an example of a Bayesian model in which it is somewhat difficult to specify a vague prior that would let the data and the collateral information speak for themselves.

The major features of the new approach are:

(1) the use of parameters with constant values for slopes across groups whenever the prior or the data indicate that this is desirable,

(2) the use of residual variance constant across groups,

(3) independent priors for regression parameters,

(4) scaling of the predictors at the grand mean rather than at the so-called "ideal scaling points" mentioned in Novick et al. (1972),

(5) transformation of all variables, including the criterion, to mean zero and variance one at the start of the calculations, with return to the raw scaling only for display of results to the user or for questions to the user.

Section 2 of this report gives a description of the old model (used by Novick et al. 1972) and a schematic comparison with the new model. Section 3 describes the old iterative algorithm for obtaining modal posterior estimates and its subsections 3a, 3b, and 3c deal with the deficiencies of that algorithm. Section 4 with its subsections 4a through 4e discusses the revisions on which the new model is based. Section 5 then outlines the new

model and derives the corresponding equations. Section 6, which is as far as possible independent of the preceding material, contains some information for users of the m-group regression program, and a final section 7 discusses possible future extensions.

2. Old model specification

In the model used by Jones and Novick (1972) and by Novick et al. (1972) for simultaneous regression in m groups, a first stage describes how the criterion is distributed given the regression parameters and given the predictor values. Considering the groups as exchangeable, the next stage treats the regression parameters (including intercept and residual variance) as a random sample from some distribution, characterized by unknown hyperparameters. A third stage specifies some, rather vague, information on these hyperparameters (c.f. Lindley and Smith, 1972).

The stages are described in Novick et al. (1972) and summarized below side by side with the new model which will be discussed in sections 4 and 5. In both models the data for the j-th individual out of the n_i individuals in the i-th group ($i = 1, 2, \ldots, m$) consist of a criterion score y_{ij} and scores on ℓ predictors x_{kij} ($k = 1, 2, \ldots, \ell$; $j = 1, 2, \ldots, n_i$). In each $(\ell + 1) \times n_i$ matrix X_i' of predictor scores we include a row of ones for the intercept. For the new model the index set $(0, 1, \ldots, \ell)$ is partitioned into two disjoint subsets F (parameters common to all groups) and G (parameters different across groups).

TABLE 1

OLD MODEL	NEW MODEL
First stage:	
$y_{ij} \sim N(\sum_{k=0}^{\ell} \beta_{ki} x_{kij};\ \phi_i)$	$y_{ij} \sim N(\sum_{f \in F} \beta_f x_{fij} + \sum_{g \in G} \beta_{gi} x_{gij};\ \phi)$
Second stage:	
$(\beta_{0i}, \beta_{1i}, \ldots \beta_{\ell i}) \sim N(\mu, H^{-1})$	$\beta_f \sim$ uniform $(-\infty, \infty)$;
$\phi_i \sim \chi^{-2}(\nu, \nu\sigma^2)$	$\beta_{gi} \sim N(\mu_g, \psi_g)$;
	$\log \phi \sim$ uniform $(-\infty, \infty)$;
Third stage:	
μ_k, ν and $\log \sigma^2 \sim$ uniform $(-\infty, \infty)$;	$\mu_g \sim$ uniform $(-\infty, \infty)$;
$H \sim$ Wishart $(\nu', \Sigma, \ell + 1)$;	$\psi_g \sim \chi^{-2}(\nu', \nu'\tau_g)$.
Σ diagonal matrix.	
User should supply (see below):	
ν' (small)	ν' (small)
diagonal elements of Σ	τ_g

This schematic presentation is restricted to the essentials. Independence assumptions, conditionings and ranges of indices are described more fully in Novick et al. (1972) for the old model and in sections 4 and 5 for the new one. For the old model, Lindley (1970)

details how integration over the hyperparameters leads to a posterior density for the regression parameters given the data. Up to an additive constant, its logarithm is (Lindley, 1970, formula 11):

$$\log p(\{\beta_{ki}\},\{\phi_i\}) =$$

$$-\sum_i (\tfrac{1}{2}n_i + 1) \log \phi_i - \tfrac{1}{2} \sum_{ij} (y_{ij} - \sum_h \beta_{hi} x_{hij})^2/\phi_i \qquad (1)$$

$$-\tfrac{1}{2}(\nu' + m - 1) \log \left| \nu'\sigma_{hk} + \sum_i (\beta_{hi} - \beta_{h.})(\beta_{ki} - \beta_{k.}) \right|$$

$$-\tfrac{1}{2}(m + 1) \log \log \{\eta(\theta^{-1} + \kappa)\}.$$

Here θ and η denote the harmonic and geometric mean of the set $\{\phi_i\}$, respectively, $\beta_{h.}$ denotes the mean across i of β_{hi} and $|a_{hk}|$, say, denotes the determinant of an $(\ell + 1) \times (\ell + 1)$ matrix A with elements a_{hk}. The constant κ is introduced to insure convergence (Lindley, 1970, page 3). For ℓ predictors and m groups, (1) is a function of $(\ell + 2)$ m parameters. Its maximization leads to the desired posterior modal estimates, but it poses some problems.

3. Problems of the old model

The computer programs made available by Novick et al. seek the maximum of (1) by the following iterative procedure. An initial set of estimates should be computed first; one might take the least squares estimates per group, the least squares estimates for the pooled sample or the so-called model II estimates, see below. Equating the derivatives of (1) with respect to β_{hi} to zero, for fixed i, leads to a set of equations which are linear in β_{hi} if one temporarily considers $\beta_{h.}$ $(h = 0, 1, \ldots \ell)$, ϕ_i and the determinant as fixed. They are

successively solved for each i; after updating means and determinant this is repeated twice. Next the updated values for all β_{hi} are used to obtain new ϕ_i by equating the derivative of (1) with respect to ϕ_i to zero; such equations are linear in $1/\phi_i$ provided that η, θ and all β_{hi} are temporarily considered as fixed. This whole process is called one iteration cycle, and such cycles should be repeated until the increase per cycle of the function (1) has become negligible.

This algorithm has been used in several applications mentioned in section 1, but not without problems:

(a) very slow convergence;

(b) non-robustness against choice of prior values for ν' and σ_{hh};

(c) non-robustness against initial choice of estimates;

(d) suboptimal determination of the mean value $\beta_{h.}$ for regression parameters for which almost total regression takes place.

3a. Slow convergence

The type of very slow convergence encountered most frequently consists of a few drastic changes in the first cycles followed by a slow, decelerating and monotone movement of each β_{hi} value to its limit. As explained in detail in Molenaar (1978), insertion after every 4th, 5th, or 6th iteration of leaps (extrapolating from the past three values in a geometric series model for each parameter separately) typically reduces the total required computer time by a factor of 2 to 4; in exceptional cases a trial run of some 10 to 20 cycles including leaps could be examined, after which a change of the default

values of the leap process produces a fully satisfactory convergence. The computer-time involved in the bookkeeping of the pre-leap values is more than gained back because one efficient leap step may produce more improvement of the goal function (1) than ten or even fifty ordinary iterations.

Since publication of Molenaar (1978) the leap process underwent two simplifications. First of all residual variance estimates were nearly always very stable across iterations, and therefore no leaps are programmed for them. Secondly, after the first few iterations, both the variances of the regression parameters (across groups) and the z-scores obtained by standardization across groups of the individual parameters in each group, were also very stable across iterations. The revised algorithm, therefore, calculates leaps only for the means across groups of the regression parameters. This means that the individual values at each iteration cycle, or their differences between two cycles, need no longer be stored for calculating the geometric series ratio underlying the leap. It suffices to extrapolate at each leap the past three mean values (across groups) of each slope and the intercept (taken at the grand mean). For each of those means a value is extrapolated from the geometric series model, and the after-leap value of each individual estimate is simply obtained by translation to the new mean value. The provisions replacing an unsatisfactory geometric series leap remain as in Molenaar (1978), with the exception that the default values now are:

-- first leap after 5 cycles;

-- each leap after 4 more cycles;

-- no leap if mean stable in 3 leading digits in last cycle;

-- leap = 20* last difference if difference changed sign or is almost zero;

-- leap = 20* last difference if last difference larger or hardly smaller than preceding difference;

--stop if log posterior density stable in 5 leading digits.

The previous version also stopped iteration when all parameters were stable in a user-specified number of digits. This provision is now deleted, because it was almost never fulfilled and led to much bookkeeping and time loss.

A FORTRAN program called BR is available in which all parameters just mentioned can be manipulated, as well as a few others. For regular use in the CADA Monitor, however, it is doubtful whether a user would have the skill to gain from successful manipulation of the parameters as compared to running some extra iterations. The BASIC version of the program, called BR1,BR2 therefore fixes all parameters at the just-mentioned default values. For the exceptional case that manipulation is desired, it could be obtained by either changing the BASIC source deck or using the FORTRAN version.

3b. Prior values robustness problem

In the log posterior density given by formula (1), the quantities σ_{hk} and ν' should be provided by the user, as a description of a plausible covariance matrix for the regression parameters and an indication of the uncertainty associated with that description (smaller ν' implying greater uncertainty). In Novick et al. (1972), it was advised to take $\nu' = 1$ unless specific prior knowledge is available. It was also advised to take the off-diagonal elements σ_{hk}, $h \neq k$, equal to zero, with the proviso of scaling the predictors at the "ideal scaling points" described in Novick et al., 1972, p. 37. This is because the intercepts can only be considered independent of the slopes when the predictors are suitably scaled. The problem of prior specification is now reduced to a choice of values for the diagonal elements σ_{hh}. If the user could provide prior estimates, say τ_h, for the variances of β_{hi}, it was advised to identify these with the prior marginal modes of these variances, namely $\nu' \sigma_{hh}/(\nu' + 2)$. For $\nu' = 1$, this leads to the specification $\sigma_{hh} = 3 \tau_h$. (This point is discussed in section 4c.)

As a practical matter, even providing τ_h values could be difficult for a user without specific prior knowledge. Therefore, Novick et al. (1972) advised setting τ_h equal to the corresponding unbiased sampling theory estimates, based on the current data, for the variances of the regression parameters. The development of these model II ANOVA estimates is given by Jackson (1972). As noted by Jackson, Novick and Thayer (1971), there are two difficulties with this advice. The first is the

theoretical point that prior quantities should not be derived from the data being analyzed. When $\nu' = 1$, however, it was hoped that the precise choice of τ_h would matter little for the posterior distribution of the regression parameters. In this light, the use of model II estimates may be seen as merely a convenient shortcut. The second difficulty is a practical one: the model II estimates may sometimes be negative. In this case, it was advised to select a "small" positive value for τ_h. As with the first point, it was hoped that the precise choice would not be too important.

The robustness of the final estimates to variations in the choice of τ_h was, in fact, illustrated for a simple case (10 groups, 1 predictor) by Jackson et al. (1971, p. 140). We shall now consider an illustration chosen to show that this robustness is not always so apparent. From the 25 percent sample of the 1968 ACT data analyzed by Novick et al (1972), 12 of the 22 groups were selected (called the "12HOMO" dataset in Molenaar, 1978). Table 2 gives the modal estimates obtained for these data when different τ_h values are used. For easier comparison, the estimates for the 12 groups have been replaced by the mean and standard deviation of those 12 values for each of the regression parameters. As before, $\nu' = 1$ and $\sigma_{hh} = 3\tau_h$ were used for all estimates in the table. Moreover, the iteration process described in section 2 always used the least squares values as initial estimates for the regression parameters. The problem of choice of initial estimates is considered in detail in section 3c.

TABLE 2. Means and standard deviations (across groups) of the posterior modal estimates for the "12HOMO" dataset as a function of the prior variance estimates τ_h (h=0,1,...,4); when these are marked "II" the model II ANOVA estimates were used, which are $\tau_0=7.97*10^{-2}$, $\tau_1=3.41*10^{-4}$ and $\tau_3=2.27*10^{-4}$ whereas the negative values for τ_2 and τ_4 were replaced. The intercept β_0 as given here pertains to the "ideal scaling", see Novick et al. (1972, p. 37)

prior variances			$10^3\beta_0$	$10^4\beta_1$	$10^4\beta_2$	$10^4\beta_3$	$10^4\beta_4$	$10^3\phi$
τ_0	τ_1'τ_3	τ_2'τ_4	M(SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
II	II	10^{-7}	- 91(202)	310(68)	174(.017)	189(23)	173(.026)	399(4.3)
II	II	10^{-6}	- 91(202)	310(68)	174(.17)	189(23)	173(.26)	399(4.3)
II	II	10^{-5}	- 90(200)	310(68)	172(1.7)	190(26)	174(2.5)	398(4.3)
II	II	10^{-4}	- 85(192)	307(68)	170(17)	190(26)	179(24)	395(4.2)
10^{-3}	10^{-5}	10^{-5}	- 89(119)	298(72)	174(32)	178(41)	165(60)	402(4.3)
10^{-2}	10^{-4}	10^{-4}	- 88(123)	304(76)	174(32)	190(42)	171(67)	398(4.2)
10^{-1}	10^{-3}	10^{-3}	-101(146)	288(114)	175(84)	218(71)	166(115)	383(4.0)
10^{-2}	10^{-5}	10^{-3}	- 97(63)	279(5.3)	184(144)	203(4.7)	163(147)	400(4.4)
10^{1}	10^{3}	10^{3}	-122(252)	248(348)	162(161)	303(301)	150(200)	372(4.1)
II	II but twice τ_1	10^{-7}	- 91(189)	311(100)	174(.016)	187(18)	173(.026)	398(4.3)
LS estimates			-123(263)	247(349)	163(160)	304(301)	150(200)	442(129)
model II estimates			- 95(270)	284(70)	163(.036)	284(67)	150(.037)	441(59)
pooled estimates			- 70	300	157	141	201	497

In the first four lines of the table, the only variation is in the small positive constant replacing the ANOVA estimates of τ_2 and τ_4 which were negative. Note that multiplication of those prior variances by a factor of 10 leads to posterior modal estimates in which the standard deviations are about 10 times as large; there is little effect on the means of β_2 and β_4 or on the other parameters. In the next 3 lines of the table we have used some priors that somebody vaguely familiar with regression equations for ACT scores might have specified. Note that the data do no longer show the almost total regression of the β_2 and β_4 values previously imposed by the very small prior τ_2 and τ_4.

In the eighth line we have purposively made τ_1 and τ_3 smaller than τ_2 and τ_4: the standard deviations of the modal estimates faithfully reflect this prior pseudo-information, although the model II ANOVA estimates were supposed to tell us that the data suggest total regression for the slopes pertaining to the second and fourth predictor, not the first and third. The ninth line shows that large prior variances produce a solution very close to the LS values. The final three lines give the characteristics of the LS estimates, the model II estimates and the regression coefficients when data from all groups are pooled into one sample.

Finally, note that Table 2 contains a line marked "II but twice τ_1", in which the only change compared to the top line is doubling the value of τ_1. The fact that the a priori most probable value of just one of the slope variances now is twice as large, i.e., the standard deviation

is multiplied by 1.414, makes the standard deviation of β_{1i} 1.464 times as large, but at the same time decreases the standard deviation of $10^4\beta_{3i}$ from 23 to 18. Looking at the modal estimates of the regression parameters themselves, the prediction for the third group changes most:

$$\text{it was: } -.431 + .018\,X_1 + .017\,X_2 + .015\,X_3 + .017\,X_4;$$

$$\text{it becomes: } -.401 + .013\,X_1 + .017\,X_2 + .016\,X_3 + .017\,X_4.$$

What conclusions can be drawn from this detailed presentation? As long as the amount of variability among regression coefficients is small, the variability of the Bayesian posterior modal estimates is strongly influenced by the prior specification; it was already noted by Novick et al. (1972) that the small positive constant replacing negative model II ANOVA estimates τ_h should be chosen with some care. The means across groups, on the other hand, are rather stable in Table 2, and it should be kept in mind that a standard deviation of .026 or of 2.5 around a mean of 176 leads to almost the same prediction equations. The quality of multiple regression equations in cross-validations is remarkably stable against changes in regression weights (Dawes, 1978, Wainer, 1976), so the differences in Table 2 may after all not be disastrous. On the other hand, in many cross-validation studies Bayesian estimates are superior only by a few percent to least squares per group, so a careful prior specification remains important. We shall resume this discussion in section 6, where the revised model will be similarly examined.

3c. Almost total regression: a threat to the model

It is well known that complete equality of parameters across groups leads to problems in Bayesian simultaneous estimation (Novick, Jackson & Thayer, 1971; Lindley, 1971; Novick et al., 1972; Novick, Lewis & Jackson, 1973). By the introduction of informative priors, Lindley, Novick and others have tried to avoid the degeneracy problems. This was satisfactory in the case of the residual variances in m-group regression, discussed in section 4a. For the slopes and the intercept, however, it does not help enough. This will be illustrated first by examining the log posterior density, and then by a numerical example. The main feature of our new model, then introduced in section 4b, was motivated by the desire to get rid of the degeneracy problem.

Let us now examine the effect of almost total regression for a parameter on the log posterior density (1) which was given on page 5. It is obvious that the first line of (1) would be maximized by the least squares (LS) values. The second line is maximized by bringing the determinant as close to zero as possible. When the user has supplied some small values for $\nu'\sigma_{hh}$ this is achieved by linear dependence among the m-vectors β_h ($h = 0, 1, \ldots \ell$). Now as soon as the estimated values of β_{hi} for some h lie very close together (almost total regression), a change in their deviations from the mean $\beta_{h\cdot}$ has almost no further influence on the residual sum of squares in the first line of (1), and thus it is used to make the determinant decrease. In other words, it pays to let the $(\ell + 1)$- variate normal distribution of the β_h

degenerate into a lower-dimensional one. Although the positive value of $\nu'\sigma_{hh}$ prevents complete degeneration, the algorithm based on the old model is deficient: because of the group-by-group calculation of new $\{\beta_{hi}\}$ a change in $\beta_{hi}-\beta_{h.}$ has far more effect on the log posterior density than a change in the mean $\beta_{h.}$, and the optimal value for $\beta_{h.}$ is never found for indices h with small variance across groups.

Table 3a shows that such undesirable behavior was indeed found for the "12HOMO" dataset used before. In each block of lines of this table, the same prior specification was combined with various initial values, described in Table 3b. Note that especially in block 1 suboptimal convergence occurs for LS, LSM or MD2 initial values; the log posterior density remains at what seems to be a local maximum, and the maximizing values of β_{hi} thus obtained differ markedly from those found with PLD initial values. Although Table 3b shows that LSM and MD2 are quite different, they lead to virtually the same modal solution in both blocks of Table 3a; the solution from LS initial values is worse, and from PLD it is better. Similar results were found for other datasets than "12HOMO".

Table 3a. Comparision of log posterior density and modal Bayesian estimates at the end of the iteration process for four sets of initial estimates described in Table 3b ("12HOMO" dataset). Within each block the same prior specification is used and thus the final log post. d. and estimates should be identical, apart from rounding errors. The algorithm was programmed to stop when the criterion remained constant in five significant digits. Instead of all 12 parameter estimates per group, their mean and standard deviation are given. The intercept as given here pertains to "ideal scaling", see Novick et al. (1972, p. 37).

Block 1: $\nu' = 1$ and prior model II with τ_2 and τ_4 (negative) replaced by 10^{-7}

initial	log post.d.	$10^3\beta_0$	$10^4\beta_1$	$10^4\beta_2$	$10^4\beta_3$	$10^4\beta_4$	$10^3\phi$
		M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
LS	280.76	- 91(202)	310(68)	174(.017)	189(23)	173(.026)	399(4.3)
LSM	281.21	-103(201)	322(69)	163(.017)	203(23)	150(.026)	399(3.9)
PLD	282.02	- 62(201)	309(69)	157(.017)	181(24)	201(.025)	399(3.5)
MD2	281.08	-103(201)	321(69)	163(.017)	203(23)	150(.026)	399(4.0)

Block 2: $\nu' = 1$ and prior model II with τ_2 and τ_4 (negative) replaced by 10^{-4}

initial	log post.d.	$10^3\beta_0$	$10^4\beta_1$	$10^4\beta_2$	$10^4\beta_3$	$10^4\beta_4$	$10^3\phi$
		M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
LS	199.27	-85(192)	307(68)	170(17)	190(26)	179(24)	395(4.2)
LSM	199.85	-83(192)	307(68)	168(17)	191(26)	180(24)	395(3.8)
PLD	200.57	-89(192)	307(68)	171(17)	192(26)	175(24)	396(3.4)
MD2	199.72	-85(192)	306(68)	169(17)	191(26)	179(24)	395(3.9)

Table 3b. Four sets of initial estimates for the "12HOMO" dataset.

initial estimates	$10^3\beta_0$	$10^4\beta_1$	$10^4\beta_2$	$10^4\beta_3$	$10^4\beta_4$	$10^3\phi$
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
LS	-123(263)	247(349)	163(160)	304(301)	150(200)	442(129)
LSM	-123(0)	247(0)	163(0)	304(0)	150(0)	442(0)
PLD	- 70(0)	300(0)	157(0)	141(0)	201(0)	497(0)
MD2	- 95(270)	251(91)	163(.036)	284(67)	150(.037)	441(59)

Explanation: LS are the least squares estimates, LSM is their mean across groups, PLD the pooled estimates taking all individuals from all groups together, MD2 are the Model II ANOVA estimates.

In a trial and error procedure not reported in Table 3, we have modified the PLD set of initial values with regard to β_{2i} and β_{4i}, the two sets of parameters which are almost totally regressed in Block 1. The final means across groups for the two sets of estimates are essentially identical with the initial values thus modified. One such modification even gives a slightly larger log posterior density than that based on PLD.

Several other trials have convinced us that the sensitivity to initial values specification is not something very exceptional, and that it seems to be most pronounced when some prior variances are specified to be very small. The initial values for such a parameter then have a mean which remains almost unchanged during the iterations, even though a change could produce a higher value of the log posterior density. This is because the algorithm adapts one β_{hi} at a time: moving it away from the slope values in the other groups is immediately punished by a decrease due to the determinant in (1). Our proposal in the next section to take β_{hi} equal across groups for certain values of h is expected to bypass this undesirable property of the present algorithm.

4. Revised model assumptions

The problems and deficiencies described above have led the authors to provide a revised model, which was schematically described in Table 1. As the algorithm based on the new model is intended for the CADA Monitor and will be regularly used on medium size computers, it was decided to

introduce a few more simplifications. The subsections 4a through 4e comment on those changes; the model itself and its consequences will be described in section 5.

4a. Constant residual variance

In a theory of Bayesian m-group regression, the groups are considered to be exchangeable, but to have varying intercepts, slopes and residual variances. A strictly common value for the latter is explicitly forbidden because it would lead to divergence problems. A small constant κ is introduced in the formulae involving the geometric and harmonic mean for just this reason. When the value of κ was varied between .01 and .0001 times the harmonic mean, this had some influence on the across groups variability of the estimated residual variances; the modal estimates of slopes and intercepts, however, remained very stable.

We have no reason to believe that homoscedasticity across groups is a more, or less, realistic assumption than homoscedasticity within groups. Moreover, in all examples of Bayesian m-group regression that we have seen the coefficient of variation of the final Bayesian estimates of ϕ_i did not exceed 1 or 2 per cent. Finally, it is found both in the algebraic formulae and in the empirical results that the Bayesian estimates of β_{hi} (which are the main goal) are hardly affected at all when small or moderate differences between ϕ_i across groups are ignored.

In the model outlined in section 5, we shall thus assume that each observation has the same residual variance ϕ, which has itself a noninformative prior proportional to ϕ^{-1}. The latter assumption

could be replaced by an inverse chi square specifying prior knowledge on ϕ. The data provide us, however, with so much information on ϕ that such prior information will not be important.

In remark 6 of section 6 of Molenaar (1978) a warning was given for a perfect or almost perfect fit in at least one group. Division by an estimated residual variance of zero, or very close to zero, could of course create problems. Now that a common value across groups is used, the risk of too small values for this residual variance has become negligible, and the previous use of a lower bound PHIMIN for residual variances has not been continued.

4b. Common values in case of low variance

It has been documented in section 3c that the algorithm does not perform well as soon as some regression coefficient shows very little variance across groups. The lack of variance may be obtained because its prior estimate is very small (the actual model II estimate might be negative, in which case Jones and Novick suggest replacement by 10^{-7}). It may also happen that the values for some parameter get very close together during the iteration process, although both the prior variance estimate and the initial values do not indicate this behavior.

In both cases, a variance of less than a user-specified bound TAUMIN is a reason for replacing all values β_{hi} ($i = 1, 2, \ldots, m$) by their mean $\beta_{h.}$; it will no longer be assumed that such a parameter is distributed across groups as a component of the multivariate normal

(μ, H^{-1}) distribution mentioned in section 2, but that it has a common value β_h which has a uniform prior distribution. Because slopes and intercepts can be very different according to the scales being used, the bound TAUMIN is applied after standardization of all variables, see subsection 4e.

In the model as described below, it is assumed that the index set $\{0, 1, 2, ., \ell\}$ denoting the intercept and the ℓ predictors is subdivided into a set F (mnemonic for fixed) for which this total regression has taken place, and its complement G (mnemonic for general) for which the values across the groups are different. The predicted value for the j-th element of the i-th group can thus be written as

$$\hat{y}_{ij} = \sum_{f \in F} \beta_f x_{fij} + \sum_{g \in G} \beta_{gi} x_{gij}$$

The $(\ell + 1)$- dimensional multinormal distribution of β_{hi} $(h = 0, 1, \ldots, \ell)$ for which some components have a variance very close to zero will thus be replaced by a vector of which some components are common to all groups, whereas the other components have a normal distribution of lower dimensionality. The actual effect on prediction of this replacement is negligible if the variance bound TAUMIN for admission to the index set F is kept low enough.

One full cycle of the iteration process now consists of four parts, (see also section 5):

(a) solution of $\{\beta_f \mid f \in F\}$ by LS regression of

$y_{ij} - \sum_{g \in G} \beta_{gi} x_{gij}$ on $\{x_{fij}\}$, treating $\{\beta_{gi}\}$ as known;

(b) solution of $\{\beta_{gi} | g \epsilon G\}$ by solving a system of linear equations which results from equating the derivatives of the log posterior density to zero, treating $\{\beta_f\}$ and $\{\sum_{i=1}^{m}(\beta_{gi}-\beta_{g.})^2\}$ as known;

(c) solution of ϕ treating all β_f and β_{gi} as known;

(d) check whether any index should pass from G to F.

The split of all indices into the subsets F and G essentially means that the revised model is really used as a class of models, or rather a lattice consisting of $2^{\ell+1}$ models because that is the number of partitions of $\{0, 1, \ldots, \ell\}$. An example is given in Figure 1.

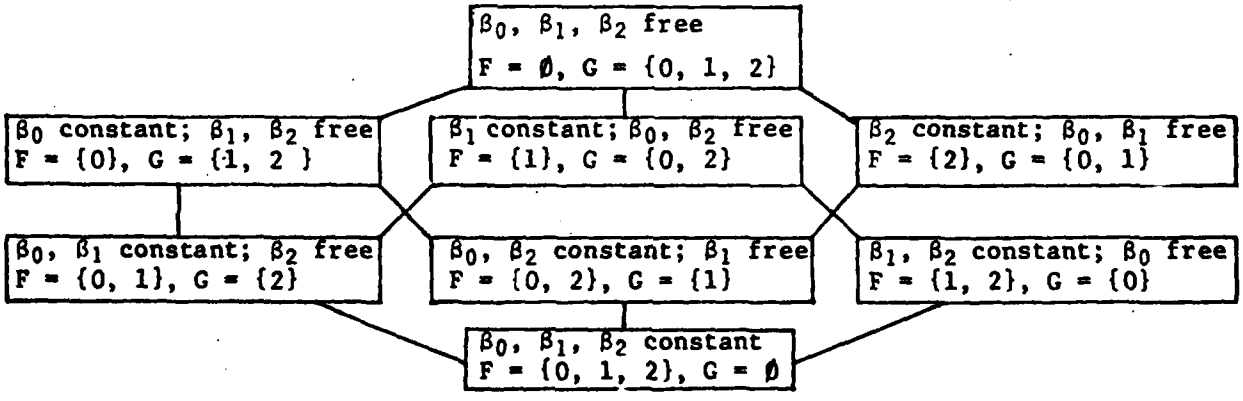

Figure 1. Lattice of 8 models for the case of $\ell = 2$ predictors.

The constant parameters at the beginning of the iterations are those with prior variance estimates less than TAUMIN, which is set at 10^{-6} in the current version. The user may force this by supplying zero entries in subjective prior estimates, or the data may force it if model II prior variance estimates are used and these come out less than

TAUMIN or even negative. During the iteration process, more indices may pass into F. The bottom model in Figure 1 with an empty G automatically produces the pooled estimates. The model used by Shigemasu (1976) is the special case with G = {0}: free intercept and constant slopes were postulated by Shigemasu, but are just one of the many possible models here.

4c. Independent priors for regression coefficients

The original model outlined in section 1 contains a multivariate normal (μ, H^{-1}) distribution for β_{hi}, and for H a Wishart $(\nu', \Sigma, \ell + 1)$ distribution. Earlier publications recommend to take $\nu' = 1$, $\sigma_{hk} = 0$ for $h \neq k$ and σ_{hh} three times a suitable prior estimate of the variance of β_{hi} (including the intercept α_i as β_{oi}).

As was explained above, the revised model allows that some β_{hi} have a common value β_h, for which a uniform prior is assumed. For the remaining parameters, say β_{gi}, it was decided to replace the assumption $\sigma_{hk} = 0$ by the slightly stronger assumption that H^{-1} itself has zero off-diagonal values. Our new model then becomes:

$$\left.\begin{array}{l} \beta_f \perp\!\!\!\perp \text{uniform } (-\infty, \infty); \\ \beta_{gi} \perp\!\!\!\perp N\ (\mu_g, \psi_g); \end{array}\right\} \text{ all } \beta_f \text{ and } \beta_{gi} \text{ independent given } \mu_g \text{ and } \psi_g$$

$$\left.\begin{array}{l} \mu_g \text{ uniform } (-\infty, \infty); \\ \psi_g \perp\!\!\!\perp \chi^{-2}\ (\nu', \nu'\tau_g) \end{array}\right\} \text{ all } \mu_g \text{ and } \psi_g \text{ independent given } \nu' \text{ and } \tau_g$$

The modification leads to a substantial simplification of the algorithm. Although prior knowledge on covariances between parameters is conceivable, it will rarely be substantial, and the revised model of course permits such covariances in the posterior distribution.

We are not advocating the use of a factor 3 in multiplying a most plausible value for the variance to find the prior value for τ_g. It was used in the old model because the mode of any $\chi^{-2}(\nu', \nu'\sigma^2)$ distribution is $\nu'\sigma^2/(\nu' + 2)$, which means for $\nu' = 1$ equating the mode to $\sigma^2/3$, and taking three times the mode for σ^2. This argument fails to take into account that the natural way to think about a variance (now called ψ_h) of a regression parameter β_h is in the logarithmic scale (that is why the uniform distribution for $\log \psi_h$ would be used as an ignorance prior). But if $\psi_h \sim \chi^{-2}(\nu', \nu'\tau_h)$ then the density of $w = \log \psi_h$ can be derived to be proportional to $\exp\{-\frac{1}{2}\nu' (w + \tau_h e^{-w})\}$ and this has its mode at $w = \log \tau_h$. An extra advantage is that the mode of the log standard deviation is now the corresponding $\log \tau_h^{\frac{1}{2}} = \frac{1}{2} \log \tau_h$. When the user is asked for a "most probable value" of the standard deviation of the true regression coefficients across groups, we prefer to use this value as a specification of $\tau_h^{\frac{1}{2}}$.

4d. Leaps for the mean only

This change has already been motivated and discussed in subsection 3a.

4e. Standardization of variables

Standardizing all predictors and the criterion to zero mean and unit variance in the pooled sample means that the predictors get the common

and absolute scale of beta weights and that the intercepts are prevented from assuming very large absolute values. This is better for the numerical accuracy, and it makes it possible to use a fixed quantity (currently 10^{-6}) for the minimum variance TAUMIN below which an index is passed to the set F and the corresponding parameter is assumed constant across groups. It is obvious that this would be undesirable for raw slopes, of which one could range from .0004 to .0008 and another from 4000 to 8000, say. At the end, just before the modal estimates are printed, a reconversion to the raw scales is made, but the intercept at the grand mean is printed as an extra column because it might be more meaningful than the intercept for all predictors zero. The desirability of the standardization was pointed out earlier in remark 6 of section 6 of Molenaar (1978).

In earlier publications by Novick et al. it was advocated to scale predictors at so called "ideal scaling points" for which the least squares estimates for the intercept and that predictor were uncorrelated across groups. Calculation of these ideal scaling points was one of the tasks of the preparatory program "BPREP" by Thayer. Our reasons for preferring the grand means, also mentioned by Novick et al. as an alternative to ideal scaling points, are the following: (a) they are easier to obtain; (b) the intercept at the grand mean is more meaningful to the user than the intercept at some "ideal point" that he never met before; (c) uncorrelatedness of the LS estimates is not the same as the (intended) uncorrelatedness of the true parameter values; and

(d) empirical evidence both from Novick and from us strongly suggests that the choice has a negligible influence on the final results.

5. The new model

The three stages of the new model have been described in Table 1, and the modified assumptions underlying it were discussed in section 4. The joint posterior density of all parameters given the data is for the new model

$$p(\{\beta_{gi}, \beta_f, \mu_g, \psi_g\}, \phi | \{x_{hij}, y_{ij}\}) \propto$$

$$\propto \phi^{-\frac{1}{2}(n+2)} \exp\left[-\frac{1}{2\phi} \sum_i \sum_j (y_{ij} - \sum_f \beta_f x_{fij} - \sum_g \beta_{gi} x_{gij})^2\right] * \qquad (2)$$

$$\prod_g \psi_g^{-\frac{1}{2}(m+\nu'+2)} \exp\left[-\frac{1}{2\psi_g}\{\nu'\tau_g + \sum_i (\beta_{gi}-\mu_g)^2\}\right];$$

here $n = \sum_i n_i$ denotes total sample size, and it is understood that in all summations i ranges from 1 to m, and j across the n_i individuals of the i-th group; moreover $f \in F$ and $g \in G$, the index sets of the constant and free parameters respectively, and all values x_{0ij} are identically 1 as dummies for the intercept.

Noting that

$$\sum_i (\beta_{gi}-\mu_g)^2 = \sum_i (\beta_{gi} - \beta_{g.})^2 + m(\beta_{g.} - \mu_g)^2, \qquad (3)$$

one integrates (2) with respect to each μ_g, and the last line of (2) becomes

$$\prod_g \psi_g^{-\frac{1}{2}(m+\nu'+1)} \exp\left[-\frac{1}{2\psi_g}\{\nu'\tau_g + \sum_i (\beta_{gi}-\beta_{g.})^2\}\right] \qquad (4)$$

Next, integration with respect to ψ_g turns this product into

$$\prod_g \{\nu'\tau_g + \sum_i (\beta_{gi} - \beta_{g.})^2\}^{-\frac{1}{2}(m+\nu'-1)}. \qquad (5)$$

The logarithm of the posterior density is thus, up to an additive constant, and omitting the dependence on the data in the left hand side:

$$\log p\ (\{\beta_{gi}, \beta_f\}, \phi) = -\tfrac{1}{2}(n+2)\log\phi +$$

$$- \frac{1}{2\phi} \sum_i \sum_j (y_{ij} - \sum_f \beta_f x_{fij} - \sum_g \beta_{gi} x_{gij})^2 +$$

$$- \tfrac{1}{2}(m + \nu' - 1) \sum_g \log \{\nu'\tau_g + \sum_i (\beta_{gi} - \beta_{g.})^2\}. \qquad (6)$$

It is instructive to compare (6) to (1). The first term is simplified because of $\phi_i = \phi$; moreover there is no final term involving geometric and harmonic means of ϕ_i. Denoting the middle term as $-\frac{1}{2\phi} Q(\beta)$, it is clear that the modal estimate for ϕ is $\tilde{\phi} = Q(\beta)/(n+2)$, and

$$\log p\ (\{\beta_{gi}, \beta_f\}, \tilde{\phi}) = -\tfrac{1}{2}(n+2)\log Q(\beta) +$$

$$+ \tfrac{1}{2}(n+2)\log(n+2) - \tfrac{1}{2}(n+2) +$$

$$- \tfrac{1}{2}(m + \nu' - 1) \sum_g \log \{\nu'\tau_g + \sum_i (\beta_{gi} - \beta_{g.})^2\}. \qquad (7)$$

This makes clear the compromise character of the modal estimates for β. The first term of (7) would be maximized by minimizing $Q(\beta)$, that is by using the least squares estimates. The last term is maximized when $\beta_{gi} = \beta_{g.}$ for each i, but when the variance is less than the bound TAUMIN, the index passes into the set F, and we would end using the pooled estimates. The point is further elaborated below.

Differentiation of (6) with respect to one fixed β_s (scF) yields

$$\sum_i \sum_j y_{ij} x_{sij} - \sum_g \sum_i \beta_{gi} \sum_j x_{gij} x_{sij} - \sum_f \beta_f \sum_i \sum_j x_{fij} x_{sij} = 0. \qquad (8)$$

If the index set F contains n_F elements and $\{\beta_{gi} | g \epsilon G\}$ are treated as known, (8) consists of n_F linear equations ($s \epsilon F$) in n_F unknowns $(\beta_f | f \epsilon F)$.

Differentiation of (6) with respect to one free parameter β_{tu} ($t \epsilon G$, $u \epsilon \{1, 2, \ldots, m\}$) yields

$$\phi^{-1}\{\sum_j y_{uj} x_{tuj} - \sum_f \beta_f \sum_j x_{fuj} x_{tuj} - \sum_g \beta_{gu} \sum_j x_{guj} x_{tuj}\} +$$

$$+ (m + \nu' - 1)(\beta_{tu} - \beta_{t.})/\{\nu'\tau_t + \sum_i (\beta_{ti} - \beta_{t.})^2\} = 0 \qquad (9)$$

Treating ϕ, $\{\beta_f | f \epsilon F\}$, $\beta_{t.}$ and the expression in the denominator as known, this is a set of $\ell + 1 - n_F$ linear equations, indexed by t, in $(\ell + 1 - n_F)$ unknown β_{gu} (for $g \epsilon G$, u fixed).

The solution for ϕ given all $\{\beta_f\}$ and $\{\beta_{gi}\}$ has already been mentioned just before (7). As announced in section 4b, each cycle of the iteration now consists of such a successive solution of all $\{\beta_f\}$ from (8), all $\{\beta_{gi}\}$ from (9) and of ϕ from $\phi = Q(\beta)/(n + 2)$. It is followed by a check, for each index $g \epsilon G$, whether $\sum_i (\beta_{gi} - \beta_{g.})^2/(m-1) <$ TAUMIN; if this is so the index passes from G to F. This check is not made after the first cycle, because the values obtained there could still be too far from the true values to justify the fixing of the parameters. Before the iterations begin, however, it is checked whether some of the prior variances (model II or user-specified) are below TAUMIN, and if so the corresponding parameters are taken constant across groups.

6. Users guide to the simultaneous regression program

This section is mainly written for the benefit of the user of the interactive m-group regression program which was a result of the research project described in this report. One may wonder why so many improved models and computer programs were produced since the publication of the basic research between 1969 and 1972. Let us try to give an indication why Bayesian simultaneous regression estimation in m-groups is a complicated matter, even compared to similar m-group models for means or proportions.

The Bayesian estimates can always be viewed as a compromise between least squares values and pooled values. Unless one of these extremes is compatible with both the data and the prior information, however, the simultaneous presence of an intercept and ℓ predictors poses an extra problem. Kelley could write $\hat{\tau}_i = \rho X_i + (1 - \rho)\, X.$, and the reliability determines the extent to which regression to the mean occurs. In our regression model, however, this extent will typically differ from parameter to parameter. Not only do we have $\ell + 1$ different extents of regression, but also each extent, and the best value to regress to, are influenced by the decisions on the other extents (cf. Jackson, 1972, p.224). And finally, when the extent was a reliability it could be estimated by one of the standard psychometric methods, but slopes and intercepts are not observable quantities, and this is an extra obstacle in trying to split their variance into true variance and error variance.

The program now used for Bayesian m-group regression has some predecessors. In 1972, the FORTRAN computer programs BPREP and BAYREG were written by Thayer and others and described in Jones and Novick (1972). A modified program MBREG, replacing BAYREG, is described in Molenaar (1978); the preparatory program MPREP proposed in that reference was never written. In the fall of 1978 MBREG was succeeded by BR, again by Molenaar, which incorporates nearly all the changes mentioned in the present report. The major exception is that BR has no rescaling of predictors and criterion. BR asks for some preprocessing of data, which could be done in the BASIC program described below, or in BPREP; independent use after different preprocessing is feasible. An input description of BR is added as Appendix A.

Lewis then turned the batch programs BPREP and BR into conversational programs in BASIC, called BR1 and BR2 respectively, and added several new features. David Chuang made some final additions, giving extra flexibility to the programs. This version will give a description of the program in that stage, reached in March 1979.

The program starts with an option of explanatory text, describing that it leads to joint modal estimates for regression coefficients in m similar (exchangeable) groups in cases of minimal prior knowledge. It specifies the restrictions (currently: at most 50 groups, at most 4 predictors, at least 6 observations per group) and announces the types of sufficient statistics per group that can be used for data entry.

Data entry may be completely via the keyboard, in a well documented but lengthy sequence of questions and answers. The standard option,

however, is data entry from a previously prepared file. Such a file goes in the current version under the local name "SOMDATA." It may now have been prepared in an earlier keyboard entry session, but after updating of the CADA Data Management capabilities it will be possible to create the complete input file there. At the end of data entry, either file or keyboard, facilities for input revision are offered.

The program next displays the least squares (LS) estimates per group for the intercept at zero, intercept at the pooled mean of the predictors, slopes and residual standard deviation. It is advisable to study these in some detail: it could be wise to delete a group or split the analysis into clusters of groups if the LS values indicate a strong violation of exchangeability or of homoscedasticity between groups. It should be kept in mind, however, that for small sample sizes the LS values behave rather wildly, and that the estimated residual standard deviations may differ by a factor of say 3 without making the model of equal s.d. seriously misleading.

As an extra line below the LS values of the last group, the pooled values (PLD) are displayed, which would be obtained by pooling the observations from all groups and calculating one least squares regression equation for the joint data. The Bayesian estimates that the program seeks to obtain can always be viewed as a compromise between the extreme situations of LS (groups have nothing to do with each other) and PLD (groups are samples from the same population).

As the next step the program calculates model II ANOVA estimates for the variance across groups of the regression parameters, and the corresponding standard deviations. The calculation, described in Jackson (1972, p. 223-224) amounts to subtracting from the "observed" variance of the LS estimates the "error" variance that can be ascribed to sampling error. It is well known that such estimates can be negative, in which case the program replaces them by zero.

For the intercept this part of the program assumes all predictors at the grand mean, which is shown on the same display. It is obvious that the intercept with all the predictors at zero could exhibit much more variability. Criterion values for the predictors at the grand mean should be more meaningful for the user, and their variability is to a large extent independent of variability in the slopes.

At this stage the user has an important option: he may delete some predictor (which may avoid multicollinearity problems) or some group (which may avoid violations of exchangeability and/or homoscedasticity).

When a satisfactory set of predictors and groups has been selected, the program proceeds to specification of prior information. This requires first prior estimates of the standard deviations across groups of the regression parameters. The user may choose either the model II estimates or provide his own prior information. In the absence of such information the model II values are certainly useful, although they have the properties of (a) making the prior data dependent and (b) ascribing all variance to sampling error whenever the estimate comes out negative, thus forcing the corresponding slope or intercept to be constant across groups. Our personal feeling is that there are

situations in which the user has no idea about true between group variability (then use model II) and also situations in which previous experience with similar regression problems enables the user to guess, at least accurately up to a factor of three, say, the prior standard deviation between groups.

When the user is doubtful as to whether these prior standard deviations are not just pure guesswork, we have two consolations for him. First, the model does not use this prior value as such, but it assumes that the true prior variance has an inverse chi square distribution with low degrees of freedom around the square of the supplied value as a typical one, so all kinds of smaller and larger variances remain possible. These degrees of freedom are the next question asked by the program: the recommended range is 1 through 10, with many groups a little higher than with few groups. For most cases df=5 will be a reasonable choice. Secondly, the user may rerun his analysis with different prior s.d. or df and find out for himself whether his results are very sensitive to his subjective decisions (our experience is that they typically are not essentially influenced unless rather little amounts of data are used.) Note that the final values of log posterior density are not comparable between runs with different prior s.d. or df.

A last choice that the user may make is whether he wants the iterations to start from LS or PLD initial estimates. It is advised to use PLD, and LS only in cases where large datasets make it plausible that the end results will be close to LS. This option is useful when the existence of bimodality is feared: if convergence from both extreme initial situations leads to the same log posterior density (up to 4 significant digits) and the same slopes and intercepts at the grand mean (up to 2 significant digits) the risks of obtaining a local maximum

are highly reduced. If the user reruns the program, after obtaining Bayesian estimates, with different prior s.d. or df, it is also possible to use the earlier Bayesian estimates as initial values. This option usually leads to faster convergence than PLD or LS initial estimates.

Now--at last--the program has enough information to start the iterative process. In each cycle several systems of linear equations have to be solved and the corresponding sets of parameters are updated. As this may be time-consuming on a medium-sized or small computer, the value of the log posterior density at the end of each cycle is printed so that the user may follow the search for its maximum. After the fifth cycle and then after each fourth next cycle there may be more increase of the log posterior density because an extrapolation or leap is made. The iteration stops when the log posterior density is stable in five significant digits. If this takes more than 10 cycles, the user may exit the iteration process after each set of 10 cycles. This facility could be useful when a restart with other initial estimates or prior values is desired. The use of the estimates obtained before stabilization of the log posterior density should not be encouraged: it is a very flat surface as a function of its many parameters, and small changes in the log posterior density may correpsond to substantial changes in the slopes and intercepts.

The next display shows the modal posterior values of intercept at zero, slopes, and intercept at grand mean. This is done for all groups, or for 10 groups at a time if there are more the 10. At the bottom the modal estimate of the residual variance and the corresponding standard deviation are given (homoscedasticity is assumed both within and between groups).

It is obvious that the user will want to keep the final modal estimates. In many cases he will be also interested to keep the prior s.d.'s and df and the LS and PLD estimates. The program therefore opens a local file DATB, in which these quantities are entered in fixed format for later use. See Appendix B for a full description. This file should be printed or copied before the next run of the program, because that run would overwrite it.

7. Conclusion, possible extensions

The new feature of this program allowing constant parameters across groups upon suggestion of either the user or the data seems to be a satisfactory solution to the problems of almost-degeneracy encountered before. Together with the extrapolation of iterations by leaps, it permits a fast and stable iterative estimation of the many parameters involved in simultaneous multiple regression. The results remain somewhat sensitive, however, to different prior specifications. Research on prior elicitation now going on in both Pittsburgh and Iowa City, may assist future users on this point.

The revised program and model are now ready for application, but the authors cannot resist the temptation to mention a few possible improvements.

Prior knowledge on means. The assumptions of uniform distributions for the parameters β_f common to all groups and for the means μ_g of parameters β_{gi} different per group could be relaxed to allow the use of prior information.

Angles for slopes. Normal distributions for slopes are not a very realistic model unless the coefficient of variation is small. Slopes for the groups that are normally distributed with e.g. a mean of 3 and a standard deviation of .2 are acceptable, but not slopes with a mean of 3 and a standard deviation of 2: the slope change from 1 to 3 is certainly more drastic than from 3 to 5, and even more when we compare a change from -1 to 3 to a change from 3 to 7. Neither uniform priors for mean slopes nor a prior for the variance of a slope independent of the mean seem to reflect our belief about slopes. Parameterization in terms of angles rather than slopes does away with most of these problems and will be examined in future research. It is not a serious drawback that it leads to non-conjugate distributions. As Bayesian modal estimates thus far have typically shown small standard deviations, the practical impact of using angles for slopes will not be dramatic.

LS estimates in restricted model. The model II estimates are obtained by subtracting sampling variance from the "observed variance" of the LS estimates. Once some of them are negative and the corresponding parameters are fixed, one could recalculate LS estimates under that restriction: common values for some parameters, free values for the others. Such a set of restricted LS estimates are useful for two purposes: they would be a better set of initial estimates for the iteration, and the model II variance estimate for the still free parameters among them is a better value for the prior

variance, because the restriction of some parameters certainly affects the mean, the raw variance, and the sampling error of the others.

Appendix A

The data deck for the FORTRAN program BR consists of the following cards:

1. Identification Card (10A8)

   Col. 1-80 Identification for data

2. Parameter Card (3I4, E8.2, F5.0, 8I2, 4F5.1)

   M and NV must be read in, other parameters get default values if blank

col.	name	format	
1-4	M	I 4	number of groups($\leq$ 25)
5-8	NV	I 4	number of predictors ($\leq$ 4)
9-12	NCMX	I 4	maximum number of cycles (default = 30 is used when 0; numbers exceeding 100 are replaced by 100)
13-19	TAUMIN	E 8.2	if prior variance, or calculated variance beyond cycle 2, is less than TAUMIN, a common value across groups is used. Default = 10^{-6} is used if number read is less than 10^{-10}.
20-25	PHIMIN	F 5.0	minimum for residual variance (default = 10^{-3} if number read is less than 10^{-7}) not used in this version.
26-27	INIST	I 2	0* = LS initial values 1* = pooled initial values 2* = model II initial values 3 = read initial values, α at ideal point

* not yet available

col.	name	format	
			4 = read initial values, α at scaling point
			5 = read initial values, α at origin.
28-29	IWR	I 2	0* = no details on iteration
			1 = details are printed
30-31	INTAU	I 2	0* = model II prior variances
			1 = read prior variances
32-33	IPUN	I 2	0 = no punched output
			1 = modal estimates are punched
			(8X, 6E 12.6)
34-35	NDH	I 2	Iteration stops when log posterior density constant in NDH leading digits (default = 5).
36-37	NDB	I 2	No leaps are taken for a mean constant in NDB leading digits (default = 4).
38-39	NCI	I 2	Number of cycles preceding first leap (default = 5, but 4 is used if number read $\leq$ 4).
40-41	NCF	I 2	Number of cycles between leaps (default = 4, is used if number read $\leq$ 4).
42-46	SCH	F 5.1	Leap = SCH* last difference if difference has just changed sign or old difference almost 0 (default = 20.0).
47-51	DCN	F 5.1	Leap = DCN* last difference if this difference is not substantially closer to 0 than previous difference (default = 20.0).

*Not yet available.

col.	name	format	
52-56	VGT	F 5.1	Not used in this version.
57-61	PNU	F 5.1	Degrees of freedom for prior variances (default = 1 is used when number read is less than 10^{-7}).

3. Prior Variance Estimates Card (6E12.6)

   Col. 1-12 τ_0 = variance estimate for intercept (ideal scaling)

   13-24 τ_1 = variance estimate for coefficient of first predictor

   ... (similarly for other predictors)

The remaining cards are read from a local file "DATA", not from INPUT, as they will remain the same for various analyses of the same dataset.

4. Predictor Card (4A8)

   Col. 1-8 Name of 1st predictor

   9-16 Name of 2nd predictor

   . .

   . .

   . .

5. Scaling Card for Original Scaling Points (5F8.0)

   Col. 1-8 Value to which criterion has been scaled

   9-16 Value to which predictor 1 has been scaled

   17-24 " " " " 2 " " "

   . .

   . .

   . .

6. Scaling Card for Ideal Points (5E13.6)

Col.	1-13	Value to which criterion has been scaled
	14-26	Ideal scaling point for predictor 1
	27-39	" " " " " 2
	.	.
	.	.
	.	.

7. Format Card for SCP Matrix (A8)

The cross products must be read in floating point form.

8. SCP Matrix Cards

For each group, there must be an upper triangular cross-product matrix punched according to the format specified by card 6. The cross-product matrices have the following form for the case of two predictors:

$$\begin{array}{l} \text{Row 1} \\ \text{Row 2} \\ \text{Row 3} \\ \text{Row 4} \end{array} \begin{bmatrix} n_i & \Sigma x_{i1j} & \Sigma x_{i2j} & \Sigma y_{ij} \\ \Sigma x_{i1j}^2 & \Sigma x_{i1j}x_{i2j} & \Sigma x_{i1j}y_{ij} & \\ \Sigma x_{i2j}^2 & \Sigma x_{i2j}y_{ij} & & \\ \Sigma y_{ij}^2 & & & \end{bmatrix}$$

These cross products are scaled to the values given by card 5.

9. Initial Values Cards (6E12.6)

For each group there must be a set of initial values, either produced by BPREP or obtained separately. For the i^{th} group, we have

Col. 1-12	initial value for β_{0i}
13-24	initial value for β_{1i}
.	.
.	.
.	.
	initial value for ϕ (must appear as the last entry on each card).

As mentioned in the text, one possible source of the information required in items 3 (Prior Variance Estimates), 6 (Ideal Scaling Points), and 9 (Initial Values), is the FORTRAN program BPREP. The information required to run that program is given by Jones and Novick (1972, p. 24).

The program BR makes use of the IMSL library routine LEQT 1F for linear equations. This routine, or a similar one, should thus be available during execution, as should be the local file "DATA" containing items 4 through 9 listed above.

Appendix B

Description of the local file "DATB" on which the program writes results important to the user:

There are at least three blocks of information. Each block consists of at least one title line, followed by m lines of numbers. These are group number, intercept at zero, intercept at pooled mean, and slopes for each of the predictors. The FORTRAN format for each of these lines is (I3, 3X, K (F10.4)), where K is the number of predictors plus two. Blocks are separated by a blank line.

The titles for the blocks are

1. PER GROUP LEAST SQUARES REGRESSION WEIGHTS

   GROUP INT(0) INT(PM) (predictor names).

2. POOLED LEAST SQUARES REGRESSION WEIGHTS.

3. BAYESIAN MODAL REGRESSION WEIGHTS

   PRIOR DF (value of ν') — PRIOR SD INT(PM) (predictor names) (values of τ_g)

4. Same as 3, for each additional Bayesian analysis after the first.

Acknowledgement

Melvin R. Novick and his staff have been extremely helpful in creating the environment in which this research could grow. David Chuang has provided useful suggestions for improvement of this paper and the computer program. We are also grateful to the secretaries in both Groningen and Iowa City who quickly and skillfully processed several versions of this report.

References

Dawes, R.M. The robust beauty of improper linear models in decision making. American Psychologist, submitted 1978.

Jackson, P.H., Novick, M.R. and Thayer, D.T. Estimating regressions in m-groups. Br. J. Math. Statist. Psychol., 1971, 24, 129-153.

Jackson, P.H. Simple approximations in the estimation of many parameters. Br. J. Math. Statist. Psychol., 1972, 25, 213-228.

Jansen, G.G.H. An application of Bayesian Statistical Methods to a Problem in Educational Measurement. Thesis Groningen University, 1977.

Jones, P.K. and Novick, M.R. Implementation of a Bayesian system for prediction in m-groups. ACT Technical Report No. 6. Iowa City, Iowa: The American College Testing Program, 1972.

Lindley, D.V. A Bayesian solution for some educational prediction problems, II. Research Bulletin 69-91, Princeton, New Jersey: Educational Testing Service, 1969.

Lindley, D.V. A Bayesian solution for some educational prediction problems, III. Research Bulletin 70-33, Princeton, New Jersey: Educational Testing Service, 1970.

Lindley, D.V. and Smith, A.F.M. Bayesian estimates for the linear model. Journal of the Royal Statistical Society (Series B), 1972, 34, 1-41.

Lindley, D.V. The estimation of many parameters. In V. P. Godambe and D.A. Sprott (Eds.) Foundations of statistical inference. Toronto: Holt, Rinehart and Winston, 1971, 435-455.

Lissitz, R.W. and Schoenfeldt, L.F. Moderator subgroups for the estimation of educational performance: a comparison of prediction models. American Educational Research Journal, 1974, 11, 63-75, followed by debate, 76-92.

Molenaar, I.W. A fast solution of the Lindley equations for the m-group regression problem. Technical Report 78-3, Iowa City, Iowa: Iowa Testing Programs, University of Iowa, 1978.

Molenaar, I.W., & Lewis, C. Bayesian m-group Regression: a survey and an improved model, Methoden en Data Nieuwsbrief SWS/Vereniging voor Statistiek, 1979, 4 (1), 62-72, Rotterdam. Netherlands.

Novick, M.R., Jackson, P.H., Thayer, D.T. and Cole, N.S. Applications of Bayesian methods to the prediction of educational performance. ACT Research Report No. 42, Iowa City, Iowa: The American College Testing Program, 1971.

Novick, M.R., Jackson, P.H., Thayer, D.T. and Cole, N.S. Estimating multiple regression in m-groups: a cross-validation study. Br. J. Math. Statist. Psychol., 1972, 25, 33-50.

Novick, M.R., Jackson, P.H., Thayer, D.T., Bayesian inference and the classical test theory model: reliability and true scores. Psychometrika, 1971, 36, 261-255.

Novick, M.R., Lewis, C. and Jackson, P.H. The estimation of proportions in m groups. Psychometrika, 1973, 38, 19-46.

Novick, M.R., Isaacs, G.L. and DeKeyrel, D.F. Computer Assisted Data Analysis - 1977. The University of Iowa.

Shigemasu, K. Development and validation of a simplified m-group regression model. Journal of Educational Statistics, 1976, 1, 157-180.

Wainer, H. Estimating coefficients in linear models: it don't make no never mind. Psychological Bulletin, 1976, 83, 213-217.

Wainer, H. On the sensitivity of regression and regressors. Psychological Bulletin, 1978, 85, 267-273.

Navy

1 Dr. Ed Aiken
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Jack R. Borsting
Provost & Academic Dean
U.S. Naval Postgraduate School
Monterey, CA 93940

1 Dr. Robert Breaux
Code N-71
NAVTRAEQUIPCEN
Orlando, FL 32813

1 MR. MAURICE CALLAHAN
Pers 23a
Bureau of Naval Personnel
Washington, DC 20370

1 Dr. Richard Elster
Department of Administrative Sciences
Naval Postgraduate School
Monterey, CA 93940

1 DR. PAT FEDERICO
NAVY PERSONNEL R&D CENTER
SAN DIEGO, CA 92152

1 Dr. Paul Foley
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. John Ford
Navy Personnel R&D Center
San Diego, CA 92152

1 CAPT. D.M. GRAGG, MC, USN
HEAD, SECTION ON MEDICAL EDUCATION
UNIFORMED SERVICES UNIV. OF THE
HEALTH SCIENCES
6917 ARLINGTON ROAD
BETHESDA, MD 20014

1 Dr. Norman J. Kerr
Chief of Naval Technical Training
Naval Air Station Memphis (75)
Millington, TN 38054

Navy

1 Dr. Leonard Kroeker
Navy Personnel R&D Center
San Diego, CA 92152

1 CHAIRMAN, LEADERSHIP & LAW DEPT.
DIV. OF PROFESSIONAL DEVELOPMMENT
U.S. NAVAL ACADEMYY
ANNAPOLIS, MD 21402

1 Dr. William L. Maloy
Principal Civilian Advisor for
Education and Training
Naval Training Command, Code 00A
Pensacola, FL 32508

1 CAPT Richard L. Martin
USS Francis Marion (LPA-249)
FPO New York, NY 09501

1 Dr. James McBride
Code 301
Navy Personnel R&D Center
San Diego, CA 92152

1 DR. WILLIAM MONTAGUE
LRDC
UNIVERSITY OF PITTSBURGH
3939 O'HARA STREET
PITTSBURGH, PA 15213

1 Library
Navy Personnel R&D Center
San Diego, CA 92152

6 Commanding Officer
Naval Research Laboratory
Code 2627
Washington, DC 20390

1 OFFICE OF CIVILIAN PERSONNEL
(CODE 26)
DEPT. OF THE NAVY
WASHINGTON, DC 20390

1 JOHN OLSEN
CHIEF OF NAVAL EDUCATION &
TRAINING SUPPORT
PENSACOLA, FL 32509

Navy

1 Psychologist
ONR Branch Office
495 Summer Street
Boston, MA 02210

1 Psychologist
ONR Branch Office
536 S. Clark Street
Chicago, IL 60605

1 Office of Naval Research
Code 200
Arlington, VA 22217

1 Code 436
Office of Naval Research
Arlington, VA 22217

1 Office of Naval Research
Code 437
800 N. Quincy SStreet
Arlington, VA 22217

5 Personnel & Training Research Programs
(Code 458)
Office of Naval Research
Arlington, VA 22217

1 Psychologist
OFFICE OF NAVAL RESEARCH BRANCH
223 OLD MARYLEBONE ROAD
LONDON, NW, 15TH ENGLAND

1 Psychologist
ONR Branch Office
1030 East Green Street
Pasadena, CA 91101

1 Scientific Director
Office of Naval Research
Scientific Liaison Group/Tokyo
American Embassy
APO San Francisco, CA 96503

1 Office of the Chief of Naval Operations
Research, Development, and Studies Branc
(OP-102)
Washington, DC 20350

Navy

1 Scientific Advisor to the Chief of
Naval Personnel (Pers-Or)
Naval Bureau of Personnel
Room 4410, Arlington Annex
Washington, DC 20370

1 LT Frank C. Petho, MSC, USNR (Ph.D)
Code L51
Naval Aerospace Medical Research Laborat
Pensacola, FL 32508

1 DR. RICHARD A. POLLAK
ACADEMIC COMPUTING CENTER
U.S. NAVAL ACADEMY
ANNAPOLIS, MD 21402

1 Mr. Arnold Rubenstein
Naval Personnel Support Technology
Naval Material Command (08T244)
Room 1044, Crystal Plaza #5
2221 Jefferson Davis Highway
Arlington, VA 20360

1 A. A. SJOHOLM
TECH. SUPPORT, CODE 201
NAVY PERSONNEL R& D CENTER
SAN DIEGO, CA 92152

1 Mr. Robert Smith
Office of Chief of Naval Operations
OP-987E
Washington, DC 20350

1 Dr. Alfred F. Smode
Training Analysis & Evaluation Group
(TAEG)
Dept. of the Navy
Orlando, FL 32813

1 Dr. Richard Sorensen
Navy Personnel R&D Center
San Diego, CA 92152

1 CDR Charles J. Theisen, JR. MSC, USN
Head Human Factors Engineering Div.
Naval Air Development Center
Warminster, PA 18974

Navy

1 W. Gary Thomson
Naval Ocean Systems Center
Code 7132
San Diego, CA 92152

1 Dr. Ronald Weitzman
Department of Administrative Sciences
U. S. Naval Postgraduate School
Monterey, CA 93940

1 DR. MARTIN F. WISKOFF
NAVY PERSONNEL R& D CENTER
SAN DIEGO, CA 92152

Army

1 Technical Director
U. S. Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 HQ USAREUE & 7th Army
CDCSOPS
USAAREUE Director of GED
APO New York 09403

1 DR. RALPH DUSEK
U.S. ARMY RESEARCH INSTITUTE
5001 EISENHOWER AVENUE
ALEXANDRIA, VA 22333

1 Dr. Myron Fischl
U.S. Army Research Institute for the
Social and Behavioral Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Beatrice J. Farr
Army Research Institute (PERI-OK)
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Milt Maier
U.S. ARMY RESEARCH INSTITUTE
5001 EISENHOWER AVENUE
ALEXANDRIA, VA 22333

1 Dr. Harold F. O'Neil, Jr.
ATTN: PERI-OK
5001 EISENHOWER AVENUE
ALEXANDRIA, VA 22333

1 Dr. Robert Ross
U.S. Army Research Institute for the
Social and Behavioral Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Robert Sasmor
U. S. Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

Army

1 Director, Training Development
U.S. Army Administration Center
ATTN: Dr. Sherrill
Ft. Benjamin Harrison, IN 46218

1 Dr. Frederick Steinheiser
U. S. Army Reserch Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Joseph Ward
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Air Force

1 Air Force Human Resources Lab
AFHRL/PED
Brooks AFB, TX 78235

1 Air University Library
AUL/LSE 76/443
Maxwell AFB, AL 36112

1 Dr. Philip De Leo
AFHRL/TT
Lowry AFB, CO 80230

1 DR. G. A. ECKSTRAND
AFHRL/AS
WRIGHT-PATTERSON AFB, OH 45433

1 Dr. Genevieve Haddad
Program Manager
Life Sciences Directorate
AFOSR
Bolling AFB, DC 20332

1 CDR. MERCER
CNET LIAISON OFFICER
AFHRL/FLYING TRAINING DIV.
WILLIAMS AFB, AZ 85224

1 Dr. Ross L. Morgan (AFHRL/ASR)
Wright -Patterson AFB
Ohio 45433

1 Dr. Roger Pennell
AFHRL/TT
Lowry AFB, CO 80230

1 Personnel Analysis Division
HQ USAF/DPXXA
Washington, DC 20330

1 Research Branch
AFMPC/DPMYP
Randolph AFB, TX 78148

1 Dr. Malcolm Ree
AFHRL/PED
Brooks AFB, TX 78235

Air Force

1 Dr. Marty Rockway (AFHRL/TT)
Lowry AFB
Colorado 80230

1 Jack A. Thorpe, Capt, USAF
Program Manager
Life Sciences Directorate
AFOSR
Bolling AFB, DC 20332

1 Brian K. Waters, LCOL, USAF
Air University
Maxwell AFB
Montgomery, AL 36112

Marines

1 H. William Greenup
Education Advisor (E031)
Education Center, MCDEC
Quantico, VA 22134

1 Director, Office of Manpower Utilization
HQ, Marine Corps (MPU)
BCB, Bldg. 2009
Quantico, VA 22134

1 DR. A.L. SLAFKOSKY
SCIENTIFIC ADVISOR (CODE RD-1)
HQ, U.S. MARINE CORPS
WASHINGTON, DC 20380

CoastGuard

1 Mr. Richard Lanterman
PSYCHOLOGICAL RESEARCH (G-P-1/62)
U.S. COAST GUARD HQ
WASHINGTON, DC 20590

1 Dr. Thomas Warm
U. S. Coast Guard Institute
P. O. Substation 18
Oklahoma City, OK 73169

Other DoD

12 Defense Documentation Center
Cameron Station, Bldg. 5
Alexandria, VA 22314
Attn: TC

1 Dr. Dexter Fletcher
ADVANCED RESEARCH PROJECTS AGENCY
1400 WILSON BLVD.
ARLINGTON, VA 22209

1 Dr. William Graham
Testing Directorate
MEPCOM
Ft. Sheridan, IL 60037

1 Military Assistant for Training and
Personnel Technology
Office of the Under Secretary of Defense
for Research & Engineering
Room 3D129, The Pentagon
Washington, DC 20301

1 MAJOR Wayne Sellman, USAF
Office of the Assistant Secretary
of Defense (MRA&L)
3B930 The Pentagon
Washington, DC 20301

Civil Govt

1 Dr. Susan Chipman
Basic Skills Program
National Institute of Education
1200 19th Street NW
Washington, DC 20208

1 Dr. William Gorham, Director
Personnel R&D Center
Office of Personnel Managment
1900 E Street NW
Washington, DC 20415

1 Dr. Joseph I. Lipson
Division of Science Education
Room W-638
National Science Foundation
Washington, DC 20550

1 Dr. John Mays
National Institute of Education
1200 19th Street NW
Washington, DC 20208

1 Dr. Arthur Melmed
National Intitute of Education
1200 19th Street NW
Washington, DC 20208

1 Dr. Andrew R. Molnar
Science Education Dev.
and Research
National Science Foundation
Washington, DC 20550

1 Dr. Lalitha P. Sanathanan
Environmental Impact Studies Division
Argonne National Laboratory
9700 S. Cass Avenue
Argonne, IL 60439

1 Dr. Jeffrey Schiller
National Institute of Education
1200 19th St. NW
Washington, DC 20208

Civil Govt

1 Dr. Thomas G. Sticht
Basic Skills Program
National Institute of Education
1200 19th Street NW
Washington, DC 20208

1 Dr. Vern W. Urry
Personnel R&D Center
Office of Personnel Managment
1900 E Street NW
Washington, DC 20415

1 Dr. Frank Withrow
U. S. Office of Education
400 6th Street SW
Washington, DC 20202

1 Dr. Joseph L. Young, Director
Memory & Cognitive Processes
National Science Foundation
Washington, DC 20550

Non Govt

1 Dr. Earl A. Alluisi
HQ, AFHRL (AFSC)
Brooks AFB, TX 78235

1 Dr. Erling B. Anderson
University of Copenhagen
Studiestraedt
Copenhagen
DENMARK

1 1 psychological research unit
Dept. of Defense (Army Office)
Campbell Park Offices
Canberra ACT 2600, Australia

1 Dr. Alan Baddeley
Medical Research Council
Applied Psychology Unit
15 Chaucer Road
Cambridge CB2 2EF
ENGLAND

1 Dr. Isaac Bejar
Educational Testing Service
Princeton, NJ 08450

1 Dr. Warner Birice
Streitkraefteamt
Rosenberg 5300
Bonn, West Germany D-5300

1 Dr. R. Darrel Bock
Department of Education
University of Chicago
Chicago, IL 60637

1 Dr. Nicholas A. Bond
Dept. of Psychology
Sacramento State College
600 Jay Street
Sacramento, CA 95819

1 Dr. David G. Bowers
Institute for Social Research
University of Michigan
Ann Arbor, MI 48106

Non Govt

1 Dr. Robert Brennan
American College Testing Programs
P. O. Box 168
Iowa City, IA 52240

1 DR. C. VICTOR BUNDERSON
WICAT INC.
UNIVERSITY PLAZA, SUITE 10
1160 SO. STATE ST.
OREM, UT 84057

1 Dr. John B. Carroll
Psychometric Lab
Univ. of No. Carolina
Davie Hall 013A
Chapel Hill, NC 27514

1 Charles Myers Library
Livingstone House
Livingstone Road
Stratford
London E15 2LJ
ENGLAND

1 Dr. John Chiorini
Litton-Mellonics
Box 1286
Springfield, VA 22151

1 Dr. Kenneth E. Clark
College of Arts & Sciences
University of Rochester
River Campus Station
Rochester, NY 14627

1 Dr. Norman Cliff
Dept. of Psychology
Univ. of So. California
University Park
Los Angeles, CA 90007

1 Dr. William Coffman
Iowa Testing Programs
University of Iowa
Iowa City, IA 52242

Non Govt

1 Dr. Meredith Crawford
Department of Engineering Administration
George Washington University
Suite 305
2101 L Street N. W.
Washington, DC 20037

1 Dr. Hans Cronbag
Education Research Center
University of Leyden
Boerhaavelaan 2
Leyden
The NETHERLANDS

1 Dr. Emmanuel Donchin
Department of Psychology
University of Illinois
Champaign, IL 61820

1 MAJOR I. N. EVONIC
CANADIAN FORCES PERS. APPLIED RESEARCH
1107 AVENUE ROAD
TORONTO, ONTARIO, CANADA

1 Dr. Leonard Feldt
Lindquist Center for Measurment
University of Iowa
Iowa City, IA 52242

1 Dr. Richard L. Ferguson
The American College Testing Program
P.O. Box 168
Iowa City, IA 52240

1 Dr. Victor Fields
Dept. of Psychology
Montgomery College
Rockville, MD 20850

1 Dr. Gerhardt Fischer
Liebigasse 5
Vienna 1010
Austria

1 Dr. Donald Fitzgerald
University of New England
Armidale, New South Wales 2351
AUSTRALIA

Non Govt

1 Dr. Edwin A. Fleishman
Advanced Research Resources Organ.
Suite 900
4330 East West Highway
Washington, DC 20014

1 Dr. John R. Frederiksen
Bolt Beranek & Newman
50 Moulton Street
Cambridge, MA 02138

1 DR. ROBERT GLASER
LRDC
UNIVERSITY OF PITTSBURGH
3939 O'HARA STREET
PITTSBURGH, PA 15213

1 Dr. Ross Greene
CTB/McGraw Hill
Del Monte Research Park
Monterey, CA 93940

1 Dr. Alan Gross
Center for Advanced Study in Education
City University of New York
New York, NY 10036

1 Dr. Ron Hambleton
School of Education
University of Massechusetts
Amherst, MA 01002

1 Dr. Chester Harris
School of Education
University of California
Santa Barbara, CA 93106

1 Dr. Lloyd Humphreys
Department of Psychology
University of Illinois
Champaign, IL 61820

1 Library
HumRRO/Western Division
27857 Berwick Drive
Carmel, CA 93921

Non Govt

1 Dr. Steven Hunka
Department of Education
University of Alberta
Edmonton, Alberta
CANADA

1 Dr. Earl Hunt
Dept. of Psychology
University of Washington
Seattle, WA 98105

1 Dr. Huynh Huynh
Department of Education
University of South Carolina
Columbia, SC 29208

1 Dr. Carl J. Jensema
Gallaudet College
Kendall Green
Washington, DC 20002

1 Dr. Arnold F. Kanarick
Honeywell, Inc.
2600 Ridgeway Pkwy
Minneapolis, MN 55413

1 Dr. John A. Keats
University of Newcastle
Newcastle, New South Wales
AUSTRALIA

1 Mr. Marlin Kroger
1117 Via Goleta
Palos Verdes Estates, CA 90274

1 LCOL. C.R.J. LAFLEUR
PERSONNEL APPLIED RESEARCH
NATIONAL DEFENSE HQS
101 COLONEL BY DRIVE
OTTAWA, CANADA K1A 0K2

1 Dr. Michael Levine
Department of Educational Psychology
University of Illinois
Champaign, IL 61820

Non Govt

1 Faculteit Sociale Wetenschappen
Rijksuniversiteit Groningen
Oude Boteringestraat
Groningen
NETHERLANDS

1 Dr. Robert Linn
College of Education
University of Illinois
Urbana, IL 61801

1 Dr. Frederick M. Lord
Educational Testing Service
Princeton, NJ 08540

1 Dr. Gary Marco
Educational Testing Service
Princeton, NJ 08450

1 Dr. Scott Maxwell
Department of Psychology
University of Houston
Houston, TX 77025

1 Dr. Sam Mayo
Loyola University of Chicago
Chicago, IL 60601

1 Dr. James A. Paulson
Portland State University
P.O. Box 751
Portland, OR 97207

1 MR. LUIGI PETRULLO
2431 N. EDGEWOOD STREET
ARLINGTON, VA 22207

1 DR. STEVEN M. PINE
4950 Douglas Avenue
Golden Valley, MN 55416

1 DR. DIANE M. RAMSEY-KLEE
R-K RESEARCH & SYSTEM DESIGN
3947 RIDGEMONT DRIVE
MALIBU, CA 90265

Non Govt

1 MIN. RET. M. RAUCH
P II 4
BUNDESMINISTERIUM DER VERTEIDIGUNG
POSTFACH 161
53 BONN 1, GERMANY

1 Dr. Peter B. Read
Social Science Research Council
605 Third Avenue
New York, NY 10016

1 Dr. Mark D. Reckase
Educational Psychology Dept.
University of Missouri-Columbia
12 Hill Hall
Columbia, MO 65201

1 Dr. Andrew M. Rose
American Institutes for Research
1055 Thomas Jefferson St. NW
Washington, DC 20007

1 Dr. Leonard L. Rosenbaum, Chairman
Department of Psychology
Montgomery College
Rockville, MD 20850

1 Dr. Ernst Z. Rothkopf
Bell Laboratories
600 Mountain Avenue
Murray Hill, NJ 07974

1 Dr. Donald Rubin
Educational Testing Service
Princeton, NJ 08450

1 Dr. Larry Rudner
Gallaudet College
Kendall Green
Washington, DC 20002

1 Dr. J. Ryan
Department of Education
University of South Carolina
Columbia, SC 29208

Non Govt

1 PROF. FUMIKO SAMEJIMA
DEPT. OF PSYCHOLOGY
UNIVERSITY OF TENNESSEE
KNOXVILLE, TN 37916

1 Dr. Kazao Shigemasu
University of Tohoku
Department of Educational Psychology
Kawauchi, Sendai 982
JAPAN

1 Dr. Richard Snow
School of Education
Stanford University
Stanford, CA 94305

1 Dr. Robert Sternberg
Dept. of Psychology
Yale University
Box 11A, Yale Station
New Haven, CT 06520

1 DR. PATRICK SUPPES
INSTITUTE FOR MATHEMATICAL STUDIES IN
THE SOCIAL SCIENCES
STANFORD UNIVERSITY
STANFORD, CA 94305

1 Dr. Hariharan Swaminathan
Laboratory of Psychometric and
Evaluation Research
School of Education
University of Massachusetts
Amherst, MA 01003

1 Dr. Brad Sympson
Office of Data Analysis Research
Educational Testing Service
Princeton, NJ 08541

1 Dr. Kikumi Tatsuoka
Computer Based Education Research
Laboratory
252 Engineering Research Laboratory
University of Illinois
Urbana, IL 61801

Non Govt

1 Dr. Maurice Tatsuoka
Department of Educational Psychology
University of Illinois
Champaign, IL 61801

1 Dr. David Thissen
Department of Psychology
University of Kansas
Lawrence, KS 66044

1 Dr. Robert Tsutakawa
Dept. of Statistics
University of MIssouri
Columbia, MO 65201

1 Dr. J. Uhlaner
Perceptronics, Inc.
6271 Variel Avenue
Woodland Hills, CA 91364

1 Dr. Howard Wainer
Bureau of Social SCience Research
1990 M Street, N. W.
Washington, DC 20036

1 DR. THOMAS WALLSTEN
PSYCHOMETRIC LABORATORY
DAVIE HALL 013A
UNIVERSITY OF NORTH CAROL
CHAPEL HILL, NC 27514

1 Dr. David J. Weiss
N660 Elliott Hall
University of Minnesota
75 E. River Road
Minneapolis, MN 55455

1 DR. SUSAN E. WHITELY
PSYCHOLOGY DEPARTMENT
UNIVERSITY OF KANSAS
LAWRENCE, KANSAS 66044

1 Dr. Wolfgang Wildgrube
Streitkraefteamt
Rosenberg 5300
Bonn, West Germany D-5300

Non Govt

1 Dr. J. Arthur Woodward
Department of Psychology
University of California
Los Angeles, CA 90024

1 Dr. Robert Woud
School Examination Department
University of London
66-72 Gower Street
London WCIE CEE
ENGLAND

1 Dr. Karl Zinn
Center for research on Learning
and Teaching
University of Michigan
Ann Arbor, MI 48104